Irmer, I.
Osaka J. Math.
54 (2017), 475–497

CRITICAL LEVELS AND JACOBI FIELDS IN A COMPLEX OF CYCLES

INGRID IRMER

(Received February 9, 2015, revised June 27, 2016)

Abstract

In this paper it is shown that the space of tight geodesic segments connecting any two vertices in a complex of cycles has finite, uniformly bounded dimension. The dimension is defined in terms of a discrete analogue of Jacobi fields, which are explicitly constructed and shown to give a complete description of the entire space of tight geodesics. Jacobi fields measure the extent to which geodesic stability breaks down. Unlike most finiteness properties of curve complexes, the arguments presented here do not rely on hyperbolicity, but rather on structures similar to Morse theory.

1. Introduction

Suppose S is a closed, oriented, connected surface of genus at least two. The complex of cycles, $\mathcal{C}(S,\alpha)$ is a variant of Harvey's complex of curves, where vertices represent multicurves in the primitive homology class α. A detailed definition is given in Section 2.

In Riemannian geometry, the dimension of the space of geodesic segments connecting any two points can be defined using the space of Jacobi fields. In Section 3.3 the "Jacobi fields" are defined and explicitly constructed, and the dimension of the space of geodesic segments is defined in Section 4.

Curve complexes are in general locally infinite, so there can be infinitely many geodesic arcs connecting two vertices. In order to be able to prove theorems in a locally infinite complex, the concept of tightness was introduced in [10], Section 4, and modified in [3], page 2. Section 2 defines tightness for $\mathcal{C}(S,\alpha)$. It is a classic result from [10], Corollary 6.14, that there are only finitely many tight geodesics connecting any two verticies m_1 and m_2 in the complex of curves $\mathcal{C}(S)$. In $\mathcal{C}(S,\alpha)$, it also follows from the main theorem of this paper that there are finitely many tight geodesics connecting any two vertices; however, unlike in $\mathcal{C}(S)$, this is not a consequence of hyperbolicity, and geodesics do not fellow travel in $\mathcal{C}(S,\alpha)$, as demonstrated in Figure 13 of [6].

In [6], Section 3, an algorithm for constructing geodesics was given, which will be outlined briefly in Section 2 for completeness. This paper develops the idea that the quantity called the "overlap function" used in this algorithm for constructing geodesics has strong parallels with a Morse function. Critical levels defined in Section 4 are analogues of conjugate points along geodesics in Riemannian geometry. The bounded topology of S gives a uniform bound on the number of critical levels, from which the theorem follows:

2010 Mathematics Subject Classification. Primary 58D19; Secondary 55U10.

Theorem 1. *Given any two vertices, m_1 and m_2 in $\mathcal{C}(S,\alpha)$, the space of tight geodesics connecting m_1 and m_2 has dimension less than $9\chi(S)^2 - 3\chi(S) - 3$.*

The *Torelli group* $\mathcal{T}$ of S is the subgroup of the mapping class group of S that acts trivially on $H_1(S,\mathbb{Z})$. The complex $\mathcal{C}(S,\alpha)$ is a member of a family of complexes that generalise the complex of curves to study $\mathcal{T}$. For example, in [2] to calculate cohomological properties of $\mathcal{T}$, in [1] to reprove a result of Birman-Powell about the generating set of the Torelli group of a surface with genus at least three, and in [7] to give a combinatorial description of a Torelli group invariant known as the Chillingworth class. Distances in these complexes are closely related to Seifert genera of links in 3-manifolds, [6], Section 6. The existence of the quasi-flats, to which the Jacobi fields studied in this paper are "tangent" is shown to have consequences for the Kakimizu complex, [8].

Surfaces in mapping tori realising Thurston norm[1]. The complex $\mathcal{C}(S,\alpha)$ was defined in analogy with Harvey's complex of curves, however the definition of $\mathcal{C}(S,\alpha)$ could be modified slightly to give a complex $\mathcal{C}^m(S,\alpha)$, for which each edge represents a subsurface with the same Euler characteristic. A mapping class ϕ fixing a connected component of $\mathcal{C}^m(S,\alpha)$ either has stable length zero, or there is an infinite geodesic invariant under the action of ϕ. This can be proven by using the observation that ϕ maps the unique "middle path" between two vertices, as defined in Section 2, into a middle path.

Take a mapping torus M_ϕ with monodromy ϕ and fiber S a surface of genus at least 2. Any element of $H_2(M_\phi;\mathbb{Z})$ has an embedded representative realising its Thurston norm, [12], Lemma 1. A ϕ-invariant geodesic in $\mathcal{C}^m(S,\alpha)$ represents an embedded surface F in the mapping torus with monodromy ϕ, where F realises the Thurston norm of the second homology class of which it is a representative, and intersects the fiber along a multicurve in $\alpha \in H_1(S;\mathbb{Z})$.

Applying the techniques of this paper to the complex $\mathcal{C}_m(S,\alpha)$, the deformations of ϕ-invariant geodesics defined by Jacobi fields could be understood as defining elementary moves to perform on Thurston norm realising surfaces to obtain further Thurston norm realising surfaces of the same homology class. This is not a 1-1 correspondence - two distinct ϕ-invariant geodesics could give the same surface in M_ϕ up to homotopy, and deforming a ϕ-invariant geodesic will not always give another ϕ-invariant geodesic. However, all such elementary moves are described by Jacobi fields or linear combinations thereof.

Sublevel Projection. The Masur-Minsky notion of subsurface projection is not directly applicable to many problems arising from studying $\mathcal{C}(S,\alpha)$. Questions relating to the way the Torelli group restricts to subsurfaces have already been shown to be central to understanding generating sets of the Torelli group, [11]. In Section 5 a notion analogous to subsurface projection from [10], Section 2, is defined by restricting to level sets of the overlap function, to which the "projections" are as rigid as possible. A distance formula analogous to that in [10], Theorem 6.12, follows from the finite number of critical levels and distance calculations in [6], Sections 4 and 5.

[1] This comment was motivated by a question/comment of Y. Minsky.

2. Background and Notation

A *curve c* in S is a piecewise smooth, injective map of S^1 into S that is not null homotopic. A *multicurve* is a union of pairwise disjoint curves on S. Let α be a primitive, nontrivial element of $H_1(S,\mathbb{Z})$. The *complex of cycles*, $C(S,\alpha)$, is the flag complex whose vertex set is the set of all isotopy classes of oriented multicurves in S in the primitive homology class α. There is an edge passing from m_1 to m_2 if m_1 and m_2 represent multicurves whose difference is isotopic to the oriented boundary of an embedded subsurface of S with the subsurface orientation. The subsurface need not be connected. The *distance*, $d_C(m_1,m_2)$, between m_1 and m_2 in $C(S,\alpha)$ is defined to be the usual path metric, where all edges have length one.

Remark. The assumption that edges of $C(S,\alpha)$ represent embedded, consistently oriented subsurfaces is not necessary for Theorem 1, but makes many definitions and discussions considerably simpler. For example, this assumption is necessary to guarantee the existence of tight geodesics between any pair of vertices, as discussed in [6] page 28. This assumption was not made in [6], so the reader should be careful translating results directly.

For simplicity, the same symbol will be used for vertices of $C(S,\alpha)$ and corresponding multicurves on S. Also, multicurves will regularly be confused with the image in S of a particular representative of the isotopy class.

The notation $m_1,\gamma_1,\gamma_2,\ldots,m_2$ will be used to denote a path γ connecting the vertices m_1 and m_2. The γ_i are the vertices the path passes through.

Tightness. Two multicurves m_1 and m_2 in general position are said to *fill S* if their complement in S is a union of discs.

The notion of "tightness" was first defined in [9], Section 4, in order to prove theorems in a complex that is not locally finite. According to the variant of the definition in [3], a path $c_0,c_1,\ldots,c_n$ in Harvey's complex of curves $C(S)$ was called *tight* at the index $\{i\neq 0,n\}$ if every curve on the surface S that crosses c_i also crosses some element of $c_{i-1}\cup c_{i+1}$. Informally, this definition ensures that c_i is contained within or on the boundary of the connected subspace of S filled by $c_{i-1}\cup c_{i+1}$. Recall that (for $C(S)$) any two multicurves representing vertices in $C(S)$ separated by a distance at least three automatically fill S. It therefore automatically follows from the definition that c_i is contained within or on the boundary of the connected subspace of S filled by $c_j\cup c_k$, for all $j<i$ and $k>i$.

However, for $C(S,\alpha)$, vertices separated by an arbitrarily large distance do not necessarily fill S, [6], page 27. For this reason, a path $\{\gamma_1,\ldots,\gamma_n\}$ in $C(S,[\gamma_1])$ is defined to be *tight* if, for every curve c in γ_i, every curve on the surface S that crosses c also crosses some element of $\gamma_j\cup\gamma_k$, for all $j<i$ and $k>i$. This definition then rules out the possibility that the set of curves corresponding to a subpath $\{\gamma_j,\gamma_{j+1},\ldots,\gamma_k\}$ of a tight geodesic enters a subsurface of S that the set of curves corresponding to two endpoints $\{\gamma_j,\gamma_k\}$ of the path do not enter.

From now on, all geodesic segments will be assumed to be tight.

Some background from [6] on how to construct geodesics will be briefly repeated here.

The *overlap function* $f_n : S\to\mathbb{Z}$ is a map from a null homologous set of curves, n, on S to a locally constant, upper semi continuous, integer valued function on S with minimum value zero. For any two points x and y in $S\setminus n$, $f_n(x)-f_n(y)$ is the algebraic intersection number of n with an oriented arc with starting point y and endpoint x.

The overlap function is not dependent on the choices of oriented arcs, because the algebraic intersection number of any closed loop with n is zero. It does however depend on the choice of representatives of the isotopy classes of curves. It will be assumed that the representatives of the homotopy classes are chosen so that the maximum, M, of the overlap function is as small as possible. When n does not contain homotopic curves, it is sufficient to assume that the curves in n are in general and minimal position. An important special case is when n is the difference of two homologous multicurves, $m_2 - m_1$. In this case, the quantity M will be called the *homological distance*, $\delta(m_1, m_2)$, between m_1 and m_2.

Corollary 2 (Corollary of Theorem 4 of [6]). *Let m_1 and m_2 be two multicurves corresponding to vertices of $C(S,\alpha)$. Then $d_C(m_1, m_2) = \delta(m_1, m_2)$.*

Surgery along a horizontal arc. Since both S and m_1 are oriented, if $t(m_1)$ is a tubular neighbourhood of m_1, $t(m_1)\smallsetminus m_1$ consists of two components; one of which can be said to be "to the right" of m_1 and the other "to the left" with respect to the orientation of m_1. An arc of $m_2 \cap (S \smallsetminus m_1)$ will be said to be *vertical* if, for any tubular neighbourhood $t(m_1)$ of m_1, the arc intersects one of the components of $t(m_1)\smallsetminus m_1$ to the left of m_1 and one of the components of $t(m_1)\smallsetminus m_1$ to the right of m_1. If an arc of $m_2 \cap (S \smallsetminus m_1)$ is not vertical, it will be said to be *horizontal*. A horizontal arc can be either to the left of m_1 or to the right of m_1. Let a be a horizontal arc with endpoints on a multicurve m. A tubular neighbourhood of $m \cup a$ has boundary consisting of a multicurve isotopic to m, and some other multicurve, call it $s_a(m)$. To *surger m along a* is to replace m with $s_a(m)$. When talking about surgering along an arc, the implicit assumption is that the arc is horizontal. Surgering along a horizontal arc clearly does not change the homology class of a multicurve.

The reason for calling arcs horizontal or vertical is illustrated in Figure 7. The overlap function is larger on one endpoint of a vertical arc than it is on the other, while a horizontal arc has both endpoints in the same level set. Suppose the first i vertices $\{m_1, \gamma_1, \ldots, \gamma_i\}$ of a tight geodesic segment connecting m_1 and m_2 have been found. When the overlap function of $m_2 - \gamma_i$ is restricted to m_2, the horizontal arcs represent local extrema. Informally, homotopy classes of horizontal arcs of $m_2 \cap (S \smallsetminus \gamma_i)$ represent the choices available in constructing the next vertex, γ_{i+1}, along the tight geodesic segment.

It is known that all tight paths, geodesic or otherwise, connecting m_1 to m_2 within $C(S,\alpha)$ can be constructed as follows: When m_1 and m_2 intersect, either surger m_1 along some set of horizontal arcs of $m_2 \cap (S \smallsetminus m_1)$, and/or discard a null homologous multicurve to obtain γ_1. When m_1 and m_2 do not intersect, the multicurve γ_1 is obtained by subtracting a null homologous submulticurve from m_1. Repeat with γ_1 in place of m_1 to obtain γ_2, etc. A proof can be found in [5], pages 3 and 4.

Middle paths. Let S_{max} be the subsurface of S on which the overlap function of $m_2 - m_1$ has its maximum and S_{imax} the subsurface of S on which the overlap function of $m_2 - \gamma_i$ has its maximum. Similarly for S_{min} and S_{imin}. Also let $S_{a \le f \le b}$ be the subsurface of S on which $a \le f_{m_2 - m_1} \le b$. When m_1 and m_2 intersect, the boundary of S_{max} is a union of horizontal arcs of $m_2 \cap (S \smallsetminus m_1)$ to the right of m_1 and horizontal arcs of $m_1 \cap (S \smallsetminus m_2)$ to the left of m_2. When m_1 and m_2 are disjoint, the boundary of S_{max} is a null homologous submulticurve of $m_2 - m_1$. It is not hard to check that, when m_1 and m_2 intersect, surgering m_1 along the arcs of $m_2 \cap (S \smallsetminus m_1)$ on the boundary of S_{max} gives a multicurve γ_1 and a curve $-\partial S_{max}$, where $\delta(\gamma_1, m_2) = \delta(m_1, m_2) - 1$ and the vertices γ_1 and m_1 are connected by an edge. Similarly, when m_1 and m_2 are disjoint, subtracting the boundary of S_{max} from m_1 gives a multicurve γ_1, where $\delta(\gamma_1, m_2) = \delta(m_1, m_2) - 1$ and the vertices γ_1 and m_1 are connected by an edge. Construct γ_2 in the same way, but with S_{1max} instead of S_{max} and γ_1 in place of m_1, similarly for γ_3, etc. A geodesic constructed in this way will be called a *middle path*.

Critical levels and level sets. If γ_i is a vertex on a middle path, informally, a critical level should be thought of as a value of i for which the level set $S_{M-i\leq f}$ is "different" from the previous level set $S_{M-i+1\leq f}$. By different, is meant either the topology, or the number of edges on the boundary of the level set changes. The critical levels along geodesic segments are therefore closely related to local extrema or saddles of the overlap function. When trying to make this notion precise, there are some technicalities involved, especially for paths that are not middle paths, so a somewhat different approach will be taken in Section 4.

Usually, a Morse theory is set up to compute a homology theory. It is not clear what the analogue, if any, of a homology theory might be in this case. Path construction in $C(S, \alpha)$ has a lot of similarities with tracing out the stable or unstable manifolds coming from the local extrema of the overlap function of $m_2 - m_1$. The finite dimensions of the space of geodesics might then be thought of as coming from the choices about the order in which different stable or unstable manifolds are traced out.

Labelling geodesic segments and surgeries. In this paper, surgeries will be denoted by listing the elements of a set of arcs along which a multicurve is surgered. The superscripts on the arcs in the set determine the multicurve along which the surgery is performed, and the subscripts label the elements in the set. For geodesic segments in a one parameter family, the superscripts will denote the element of the family, and the subscripts determine the vertex of a geodesic segment.

2.1. Independent Surgeries. When making statements about how to perturb the geodesic segment $m_1, \gamma_1, \ldots, m_2$, it is necessary to have a concept of what surgeries are equivalent to or dependent on each other. In order to understand this, we first need a notation for the smallest subsurface inside which a multicurve is altered by a surgery and the subsequent isotopy to put it in minimal (but not general) position with m_2.

If a is a connected component of $m_2 \cap (S \setminus m_1)$, it will be said to be *homotopic* to another arc b of $m_2 \cap (S \setminus m_1)$ if it can be homotoped onto b by a homotopy that keeps the endpoints of a on m_1.

For multicurves in minimal position, a homotopy class of arcs with representative a determines a rectangle $R(a)$ in S, as shown in Figure 1. The "short sides" of $R(a)$ are arcs in the homotopy class. When the homotopy class only has a single representative, $R(a)$ is degenerate and consists of a single arc.

Suppose γ_i is surgered along an arc a of $m_2 \cap (S \setminus \gamma_i)$, the resulting multicurve is put in minimal position with respect to m_2, and a null homologous multicurve $N(a)$ is discarded to obtain γ_{i+1}. Here, $R(a)$ represents the smallest possible subsurface through which the surgered multicurve must be moved to put it in minimal (but not general) position with respect to m_2. Alternatively, we might want to surger γ_i along an arc b of $m_2 \cap (S \setminus \gamma_i)$, and discard the null homologous curve $N(b)$. When $R(a)$ is disjoint from $R(b)$, $N(a)$ is disjoint from the long sides of $R(b)$ and $N(b)$ is disjoint from the long sides of $R(a)$, it is possible to perform either surgery, or both, independently of each other. In this case, we will say that

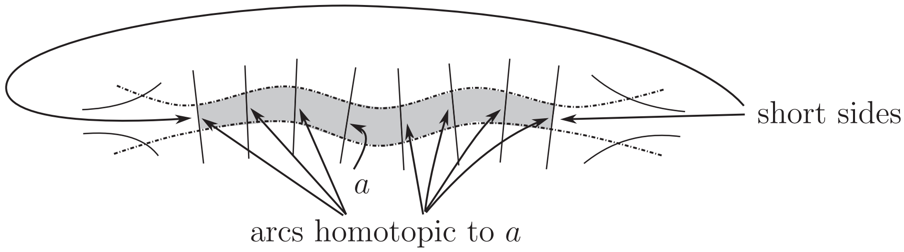

FIG. 1. The rectangle representing a homotopy class of arcs. The thin black lines are subarcs of m_2 and the striated lines are subarcs of m_1.

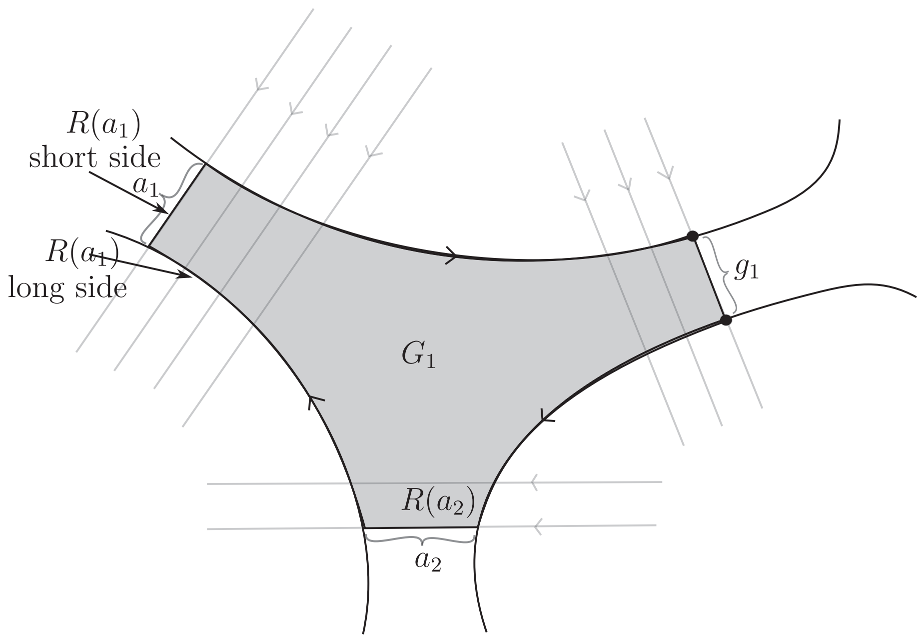

FIG. 2. The subsurface through which the multicurve γ_i must be isotoped to put it in minimal position with respect to m_2 after being surgered along the arcs $\{a_1, a_2\}$ is shaded. This subsurface contains the polygon G_1 with edges along m_2 homotopic to the arcs $\{a_1, a_2, g_1\}$. The grey lines are subarcs of m_2 and the black lines are subarcs of γ_i.

the surgery along a is *independent* of the surgery along b. If no null homologous curves are discarded, $N(a)$ and/or $N(b)$ are understood to be empty sets.

When $\{a_j^{i+1}\}$ is a collection of horizontal arcs of $m_2 \cap (S \setminus \gamma_i)$, as shown in Figure 2, the smallest subsurface through which the surgered multicurve must be isotoped to obtain a multicurve in minimal position with respect to m_2 might be larger than the union of the rectangles $\{R(a_j^{i+1})\}$. It could happen that all but one, g_k, of the arcs of $m_2 \cap (S \setminus \gamma_i)$ on the boundary of a polygon G_k in $S \setminus (m_2 - \gamma_i)$ are homotopic to one of the arcs $\{a_j^{i+1}\}$. Then the smallest subsurface through which the surgered multicurve must be isotoped to put it in minimal position with respect to m_2 is the union $\cup_j R(a_j^{i+1}) \cup_k G_k \cup_k R(g_k)$, as shown in Figure 2. In this case, the surgeries corresponding to $\{a_j^{i+1}\}$ will be said to be *independent* *of the surgeries corresponding to the arcs* $\{b_l^m\}$ if $\cup_j R(a_j^{i+1}) \cup_k G_k \cup_k R(g_k)$ is disjoint from $\cup_l R(b_l^m) \cup_n G_n \cup_n R(b_n)$, $\cup_j N(\{a_j^{i+1}\})$ is disjoint from the long sides of $\cup_l R(b_l^m) \cup_n R(b_n)$ and $\cup_j N(\{b_j^{i+1}\})$ is disjoint from the long sides of $\cup_j R(a_j^{i+1}) \cup_k R(g_k)$.

Equivalent Surgeries. It can happen that two independent surgeries, followed by discarding different null homologous submulticurves can give the same result up to isotopy. Two such surgeries will be said to be *equivalent*. An example of this can be found in Example 3. The curve γ_9 is obtained from m_2 by applying a bounding pair map four times. There are two horizontal arcs of $m_2 \cap (S \setminus \gamma_9)$, and surgering along either of them results in untwisting one pair of twists.

Throughout this paper, the notation $\{a_j^{i+1}\}$ is used to refer to a set of arcs corresponding to surgeries performed on γ_i to obtain the next vertex, γ_{i+1} of the tight geodesic segment $m_1, \gamma_1, \ldots, m_2$. Since different arcs can give rise to equivalent surgeries, this set might not be uniquely defined, hence will be referred to as a choice of arcs representing the surgeries performed on γ_i to obtain γ_{i+1}.

3. Jacobi Fields

Subsection 3.1 recalls the definition of Jacobi fields from Riemannian geometry. In order to motivate the definition of Jacobi fields for $C(S, \alpha)$, it helps to have a few examples of one parameter families in mind. These examples are given in Subsection 3.2. Subsection 3.3 then defines and constructs one parameter families and their associated Jacobi fields in $C(S\,; \alpha)$. Finally, Subsection 3.4 makes rigorous the notion of a linear combination of Jacobi fields.

3.1. Jacobi Fields in Riemannian Geometry. This material can be found in most books about Riemannian geometry, for example Chapter 5 of [4]. Informally, a Jacobi field determines the relative motion of two nearby particles in free fall in space-time. Given a smooth 1-parameter family of geodesics, γ_s, with $\gamma_0(t) := \gamma(t)$, a *Jacobi field* is a vector field along the geodesic γ that satisfies the Jacobi equation

$$\frac{D^2 J(t)}{dt} + R(J(t), \dot{\gamma(t)})\dot{\gamma(t)} = 0$$

where $\frac{D}{dt}$ is the covariant derivative with respect to the Levi-Civita connection, R is the curvature tensor, and $\dot{\gamma(t)}$ is a tangent vector field to γ depending on the parameterisation t.

Equivalently, Jacobi fields can be thought of as tangent vectors to 1-parameter families of geodesics;

$$J(t) = \left.\frac{d\gamma_s(t)}{ds}\right|_{s=0}$$

In this paper we will not be interested in "trivial" Jacobi fields tangent to γ and coming from a change in parameterisation.

A helpful example of Jacobi fields to think about is on S^2, with constant sectional curvature K. Consider a 1-parameter family of geodesic segments $\gamma_s(t)$ running from the south pole of S^2 to the north pole, where t is arc length along γ. The Jacobi equation then becomes

$$\frac{D^2 J}{dt^2} + KJ = 0$$

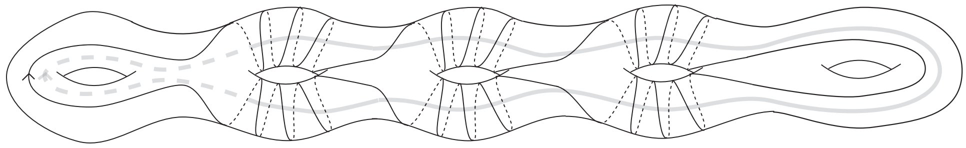

Fig. 3. The curves m_1 and m_2 (grey) from Example 3. The arc a_1 is the fat dotted grey line.

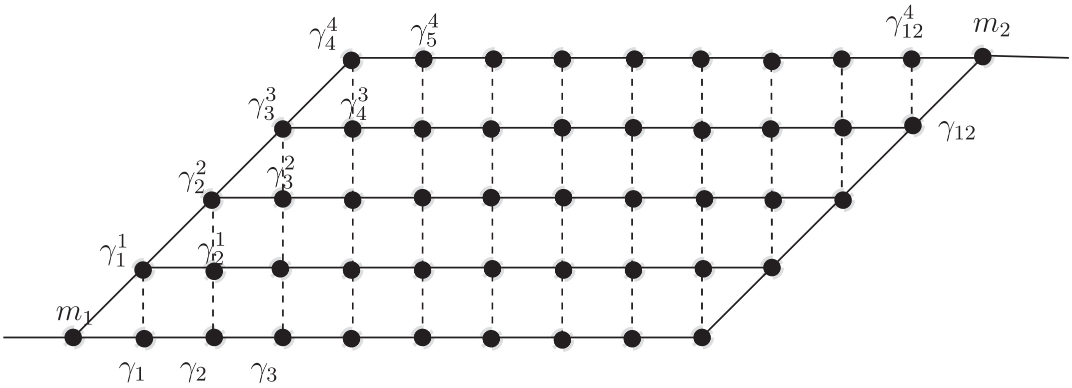

Fig. 4. A one parameter family of geodesics in $\mathcal{C}(S, [m_1])$ from Example 3. Geodesic segments in the one parameter family are represented by solid lines, other edges in the graph by dotted lines.

(This is Equation 2 Chapter 5 of [4]), a solution of which gives a Jacobi field everywhere orthogonal to γ with magnitude

$$\frac{\sin(t\sqrt{K})}{\sqrt{K}}.$$

The north and south poles are therefore *conjugate points* along γ, namely points at which a nontrivial Jacobi field along γ goes through zero.

Jacobi fields can be thought of as generators of "local isometries"; a Killing vector field, when restricted to γ, gives a Jacobi field, although the converse is not true.

3.2. Examples. This subsection gives some examples of one parameter families of geodesic segments.

Example 3 (Alternative Surgeries). The curves m_1 and m_2 are shown in Figure 3. There are two horizontal arcs of $m_2 \cap (S \setminus m_1)$. One of them is depicted in Figure 3 as a dotted grey line; denote it by a_1. The geodesic $m_1, \gamma_1, \ldots, \gamma_{12}, m_2$ is the unique middle path connecting m_1 and m_2. Recall that γ_1 is constructed by first surgering m_1 along the arc a_1 and discarding a resulting null homologous multicurve. The curve γ_1 has one fewer of the pairs of twists furthest to the left. The curve γ_2 is obtained similarly by surgering along an arc $v_1 \circ a_1 \circ v_2$, where v_1 and v_2 are arcs of $m_2 \cap (S \setminus m_1)$ to either side of a_1. This surgery undoes the next leftmost pair of twists. The curves γ_3 and γ_4 are obtained similarly. Once we get to γ_5, we start unwinding pairs of twists inside the genus one subsurface to the right of the first subsurface. Last of all, the twists inside the rightmost subsurface are undone.

The decisions involved in constructing $m_1, \gamma_1, \ldots, \gamma_{12}, m_2$ were completely arbitrary. For example, for $k \leq 4$ we could construct a family of geodesic segments $m_1, \gamma_1^k, \ldots, \gamma_{12}^k, m_2$, as follows: $m_1, \gamma_1^k, \ldots, \gamma_{12}^k, m_2$ is the geodesic segment obtained by first untwisting k twists, working from right to left, and then untwisting from left to right. This family of geodesic segments in $C(S, \alpha)$ is depicted in Figure 4.

In this example, $1 \leq k \leq 4$. If we were to try to construct γ^5 in the obvious way, in $C(S, \alpha)$ it would be no further from γ than γ^4. This is because the surgery performed on γ_4^5 to obtain γ_5^5 is equivalent to the surgery performed on γ_4 to obtain γ_5. The surgery in question undoes one of the twists in the middle. Even in this very simple example, thereare a couple of one parameter families

In Example 3, the one parameter family contained four geodesics. We can not meaningfully increase this number and still obtain a one parameter family. For example, as indicated in Figure 4, the one parameter family does *not* contain the tight geodesic segment m_1, γ_1, γ_2^1, γ_3^1, γ_4^1, γ_5^1, γ_6^1, γ_7^1, γ_8^1, γ_9^1, γ_{10}^1, γ_{11}^1, γ_{12}^1, m_2. This is because a one parameter family is required to have the property that the *set of geodesics* themselves can be parameterised by one parameter, not just the *set of points* on the set of geodesics. If the one parameter family contained this geodesic in addition to the four others, it would not be clear in what direction the parameter is increasing at the vertex γ_1 - in the direction of the vertex γ_1^1 or γ_2^1?

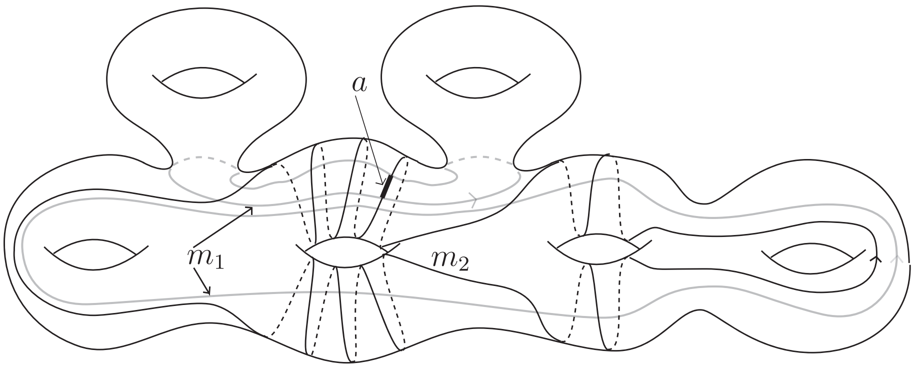

FIG. 5. The multicurve m_1 is shown in grey.

EXAMPLE 4 (Optional Surgeries). Figure 5 shows the multicurves m_1 and m_2. In this example, the geodesic $m_1, \gamma_1, \ldots, \gamma_5, m_2$ is the middle path connecting m_1 and m_2, constructed

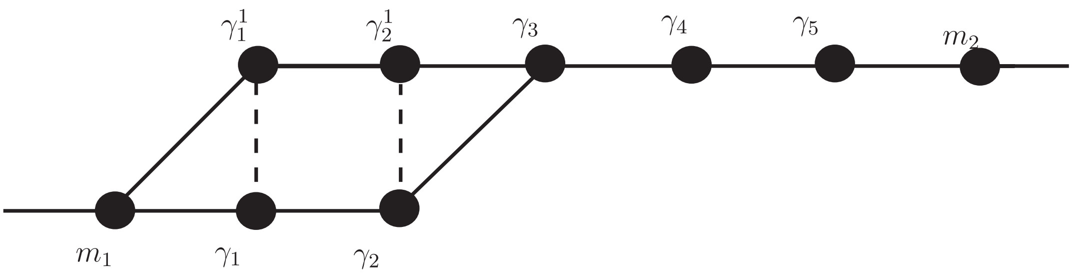

FIG. 6. The one parameter family of geodesics in $C(S, [m_1])$ from Example 4. Geodesic segments of geodesics in the one parameter family are represented by solid lines.

as outlined in Section 2. In Figure 5 this corresponds to untwisting from right to left. When constructing γ_1, in addition, we might also have surgered along the arc a shown. If we do not do this, at the very latest, γ_2 has to be surgered along a set of arcs including a to obtain γ_3, otherwise γ would not be a geodesic. In this example, for $i = 1, 2$ γ_i^1 represents a multicurve obtained from γ_i by applying s_a, and for $i = 3, 4, 5$ γ_i^1 and γ_i represent the same multicurves.

This gives a (small) one parameter family $\{\gamma, \gamma^1\}$ depicted in Figure 6.

3.3. One parameter families and Jacobi fields. In this subsection, different ways of constructing one parameter families of geodesic segments will be discussed. The main difficulty is in understanding the circumstances under which these constructions can be applied without causing contradictions with path construction on some other subsegment. The one parameter families are used to define "Jacobi fields". Before defining one parameter families, it is necessary to establish a canonical choice of representatives of isotopy classes of multicurves, so as to be able to identify arcs of $m_2 \cap (S \setminus \gamma_i)$ for different values of i.

Choices of Isotopy Classes. Put m_1 and m_2 in general and minimal position. The representative of the isotopy class γ_1 is then obtained as follows: first perform the surgeries corresponding to $\{a_i^1\}$ on m_1. Recall that $\cup_i R(a_i^1) \cup_j G_j \cup_j R(g_j)$ is the smallest subsurface through which the multicurve m_1 must be isotoped to put it in minimal position with m_2 after surgeries along $\{a_i^1\}$. The multicurve obtained by surgering, isotoping and perhaps discarding a null homologous submulticurve is called γ_1. Therefore, any subarc of γ_1 either coincides with m_1 outside of an ϵ-neighbourhood of $\cup_i R(a_i^1) \cup_j G_j \cup_j R(g_j)$, or if it has become part of a null homologous submulticurve, it might have been discarded. The multicurve γ_2 is then obtained by performing the surgeries corresponding to $\{a_i^2\}$ on this representative of the isotopy class γ_1. Any part of the resulting multicurve outside of an ϵ-neighbourhood of $\cup_i R(a_i^2) \cup_j G_j \cup_j R(g_j)$ either coincides with γ_1 or is discarded, etc.

Jacobi fields come about in a few different ways; from optional surgeries, alternative surgeries or choices about null homologous submulticurves. First of all Jacobi fields coming from optional surgeries will be defined.

Recall that the path γ is a geodesic passing through the vertices $m_1, \gamma_1, \gamma_2, \ldots, m_2$.

Optional Surgeries. Let a be a horizontal arc of $m_2 \cap (S \setminus \gamma_i)$, that defines an *optional* surgery in the following sense: s_a is independent of the surgeries along the set of arcs $\{a_j^{i+1}\}$ performed on γ_i to obtain γ_{i+1}. In addition, surgering along the set of arcs $\{a\} \cup \{a_j^{i+1}\}$ determines an edge of $\mathcal{C}(S, \alpha)$.

In Example 4, there was a subinterval $\{m_1, \gamma_1\}$ of γ along which s_a determined an optional surgery. Further along γ at vertex γ_2, s_a was one of the surgeries performed to obtain γ_3. As a result, the path γ^1 and γ then converged on vertex γ_3.

Now consider the general case in which s_a determines an optional surgery along a subinterval I_a of γ. We can try to construct γ^1 as follows: before the interval I_a is reached, $\gamma_i^1 = \gamma_i$. In the interval I_a, γ_i^1 is obtained from γ_i by applying s_a. How is γ_i^1 defined for the remaining values of i?

There is a smallest i^* such that for $i^* \leq i$, m_2 does not cross γ_i at one or both of the points identified with the endpoints of a. So if s_a is not equivalent to a surgery applied to γ_{i^*-1}, then either one or both of the endpoints of a are on a null homologous multicurve that was discarded after γ_{i^*-1} was surgered along $\{a_j^{i^*}\}$, or $\{a_j^{i^*}\}$ can be chosen to contain an arc b,

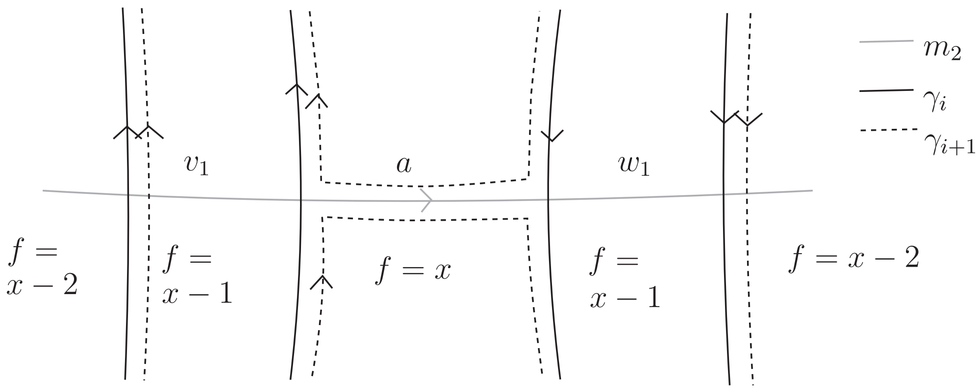

Fig. 7. Consecutive surgeries.

where b has an endpoint in common with a. The surgery s_a is then no longer defined on γ_i, for $i^* \leq i$. Tightness of γ^1 rules out the possibility of applying s_a^{-1} to an element of γ_i^1. Without destroying tightness of γ^1, for $i^* \leq i$ we can not obtain γ_i^1 from γ_i by surgering along a set of arcs of $m_2 \cap (S \setminus \gamma_i)$. In other words, for $i^* \leq i$ there is no edge in $C(S,\alpha)$ connecting γ_i^1 and γ_i, and no obvious way of defining a one parameter family that somehow relates γ_i^1 to γ_i.

In summary - the surgery s_a does not determine a one parameter family unless it is equivalent to a surgery that is actually performed somewhere along γ.

One parameter families. It will now be explained how to construct the first geodesic above γ in a one parameter family coming from an optional surgery. Let I_a be the largest subinterval of $m_1, \gamma_1, \gamma_2, \ldots, m_2$ on which the arc a determines an optional surgery on the preceding multicurve. By definition the last vertex of I_a is the vertex of γ that is surgered along the arc a, for the reasons explained in the previous paragraph. For γ_i in I_a, let γ_i^1 be obtained from γ_i by applying s_a. The vertex γ_i^1 coincides with γ_i outside of I_a. The geodesic segment $m_1, \gamma_1^1, \gamma_2^1, \ldots, m_2$ is the first element of the one parameter family above γ.

After surgering along a horizontal arc a of $m_2 \cap (S \setminus \gamma_i)$, suppose a new horizontal arc, $v_1 \circ a \circ w_1$ is created, as shown in Figure 7. Here $v_1 \circ a \circ w_1$ is the arc of $m_2 \cap (S \setminus \gamma_i^1)$ obtained by concatenating a with arcs v_1 and w_1 of $m_2 \cap (S \setminus \gamma_i)$ on either side of it. Surgering γ_{i+1} along $v_1 \circ a \circ w_1$ will be thought of as being the most obvious continuation of the surgery along a.

Denote by γ^2 the second element of the one parameter family above γ. If $v_1 \circ a \circ w_1$ is not a horizontal arc of $m_2 \cap (S \setminus \gamma_i^1)$ that represents an optional surgery for some i, the one parameter family consists of $\{\gamma, \gamma_1\}$ only. Otherwise, $m_1, \gamma_1^2, \gamma_2^2, \ldots, m_2$ is constructed from $m_1, \gamma_1^1, \gamma_2^1, \ldots, m_2$ analogously to the way $m_1, \gamma_1^1, \gamma_2^1, \ldots, m_2$ was constructed from $m_1, \gamma_1, \gamma_2, \ldots, m_2$. Let n be the natural number such that the one parameter family can not be extended past $m_1, \gamma_1^n, \gamma_2^n, \ldots, m_2$.

The one parameter family $\mathcal{P}(a,\gamma)$ of geodesics over γ consists of the geodesics $\{\gamma, \gamma^1, \ldots, \gamma^n\}$.

Jacobi Fields. A *vector* in $C(S,\alpha)$ is defined to be an oriented geodesic segment, where the magnitude of the vector is the length of the geodesic segment. A *vector field* along γ is defined to be a map from the vertices γ_i of γ into vectors, where γ_i maps to a vector with basepoint γ_i. A *Jacobi field*, $J(a,\gamma)$, is a vector field along γ, that is "tangent" to the one parameter family $\mathcal{P}(a,\gamma)$ in the following sense: $J(a,\gamma)$ restricted to the vertex γ_i of γ is the geodesic segment passing through the vertices $\gamma_i, \gamma_i^1, \gamma_i^2, \ldots, \gamma_i^n$. The *support* of $J(a,\gamma)$ is the largest subpath of $m_1, \gamma_1, \gamma_2, \ldots, m_2$ for which $\gamma_i^k \neq \gamma_i$ for some $k \leq n$.

Alternative Surgeries. The multicurve γ_{h+1} is constructed from γ_h by surgering along arcs $\{a_j^{h+1}\}$ and possibly discarding a null homologous submulticurve $N(h+1)$. Suppose that, alternatively, a geodesic segment could have been constructed by surgering γ_h along the arcs $\{c_j^{h+1}\}$ instead of a subset $\{b_j^{h+1}\}$ of $\{a_j^{h+1}\}$, and possibly discarding a null homologous submulticurve $N(c,h+1)$. Whenever the set $\{c_j^{h+1}\}$ could not have been replaced by a smaller subset, the surgeries along $\{c_j^{h+1}\}$ will be called *alternative surgeries*. It will be assumed that the surgeries along $\{c_j^{h+1}\}$ are not equivalent to the surgeries along $\{b_j^{h+1}\}$.

In this subsection, we will be interested in very specific alternative surgeries. Suppose γ is the middle path connecting m_1 and m_2. Then recall that $\{a_j^{i+1}\}$ are the arcs of $m_2 \cap (S \setminus \gamma_i)$ on ∂S_{imax}, where S_{imax} is the subsurface of S on which the overlap function of $m_2 - \gamma_i$ has its maximum. This is assuming the representative of the homology class of γ_h outlined in Subsection 3.3. Define $\{c_j^1\}$ to be the set of arcs of $m_2 \cap (S \setminus m_1)$ on ∂S_{min}.

When the surgeries along $\{c_j^1\}$ are not equivalent to the surgeries along $\{a_j^1\}$, there is a one parameter family, constructed as follows: The multicurve γ_1^1 is constructed by surgering m_1 along $\{c_j^1\}$ and discarding $-\partial S_{min}$. For $1 < i$, the vertices γ_{i+1}^1 lie along the middle path connecting γ_1^1 to m_2.

Let $\{c_j^2\}$ be the arcs of $m_2 \cap (S \setminus \gamma_1^1)$ on the boundary of the subsurface S_{1min}^1 on which the overlap function of $m_2 - \gamma_1^1$ takes on its minimum, assuming the representative of the isotopy class of γ_1^1 outlined in Subsection 3.3. When the surgeries on γ_1^1 along $\{c_j^2\}$ are not equivalent to surgeries along $\{a_j^1\}$, the next geodesic segment in the one parameter family, γ^2, is constructed as follows: γ_1^2 coincides with γ_1^1. The multicurve γ_2^2 is constructed by surgering γ_1^1 along $\{c_j^2\}$ and discarding $-\partial S_{1min}^1$. For $2 < i$, the vertices γ_{i+1}^2 lie along the middle path connecting γ_2^2 and m_2.

Further geodesic segments in the one parameter family, if any, are constructed analogously. If m_1 and m_2 are replaced by γ_j and γ_k, $j < k$, a one parameter family can be constructed above the geodesic segment $\gamma_j, \gamma_{j+1}, \ldots, \gamma_k$. This can then be extended to a one parameter family above γ by setting $\gamma_m^l := \gamma_m$ for $m \leq j$ and $k \leq m$.

For all other alternative surgeries, suppose $\{c_j^{h+1}\}$ is a set of arcs corresponding to surgeries that are actually performed somewhere along γ. To be more precise, suppose that for $h < k$, $\{a_j^{k+1}\}$ can be chosen to coincide with $\{c_j^{h+1}\}$. Since $\{c_j^{h+1}\}$ define surgeries along γ_h and γ_k, assuming the the representatives of isotopy classes as outlined in Subsection 3.3, then $\{c_j^{h+1}\}$ also define surgeries along γ_i for $h < i < k$. If for all $h < i < k$, the surgeries on γ_i along $\{c_j^{h+1}\}$ are independent of the surgeries along $\{a_j^{i+1}\}$, then these surgeries commute. A one parameter family can therefore be constucted as follows: for $i \leq h$ and $k + 1 \leq i$, γ_i^1 coincides with γ_i. The multicurve γ_{h+1}^1 is constructed by surgering γ_h along $\{c_j^{h+1}\}$ and perhaps discarding a null homologous submulticurve. For $h + 1 < i \leq k$, γ_{i+1}^1 is constructed by surgering γ_i^1 along $\{a_j^i\}$ and perhaps discarding a null homologous submulticurve.

Denote by v_j and w_j the arcs of $m_2 \cap (S \setminus \gamma_h)$ on either side of c_j^{h+1}. Suppose $\{a_j^{k+2}\}$ can be

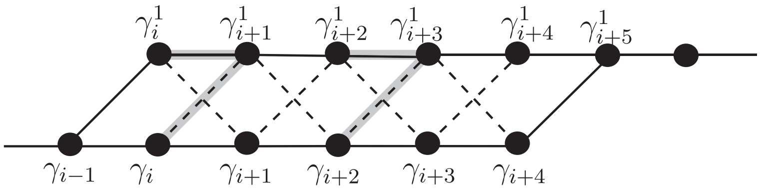

FIG. 8. When γ_i^1 and γ_i are distance two in $C(S,[m_1])$, examples of canonical geodesic paths connecting γ_i^1 and γ_i are shown in grey.

chosen to coincide with $\{v \circ c_j^{h+1} \circ w\}$, and $\{v \circ c_j^{h+1} \circ w\}$ is independent of all the surgeries along $\{a_j^{i+1}\}$, for $h+1 < i < k+1$. Then it is possible to construct a second geodesic in the one parameter family. This construction is analogous to the construction of γ^1 from γ. Similarly for further geodesic segments in the one parameter family, if any.

REMARK 5. For alternative surgeries, two constructions of one parameter families have been given. In both of these cases, it is possible that γ_i^1 is distance two from γ_i in $C(S,[m_1])$ for some values of i. However, γ^1 is still Hausdorff distance one from γ in $C(S,[m_1])$, since by construction, γ_{i+1}^1 is at most distance one from γ_i. It is therefore still possible to define the canonical geodesic path in $C(S,[m_1])$ between γ_i^1 and γ_i, required in the definition of Jacobi field. This is illustrated in Figure 8.

Null homologous submulticurves. Suppose γ_i has null homologous submulticurves $N_1, N_2, \dots N_m$ in the oriented isotopy class n. Assume that n bounds an embedded subsurface, because otherwise discarding a submulticurve isotopic to n does not determine an edge. It can happen that discarding one of these submulticurves from γ_i decreases the homological distance from m_2. When this happens, n will be called *nonperipheral* in γ_i, otherwise n is *peripheral* in γ_i. In Example 4, the multicurve m_1 has a peripheral null homologous submulticurve.

When the subsurface bounded by the perpheral null homologous multicurve N_m is disjoint from the subsurface bounded by $\gamma_{i+1} - \gamma_i$, discarding N_m is optional, and one parameter families can be constructed as for optional surgeries. A one parameter family is obtained when, for some $i < k$, the null homologous submulticurve N_m is discarded from γ_k. A second geodesic segment in the one parameter family is constructed by taking the most obvious continuation of discarding N_m, namely discarding N_{m-1}, etc.

When n is nonperipheral, discarding N_m is analogous to an alternative surgery. Discarding N_m commutes with any set of surgeries along arcs whose endpoints are not on N_m, and it is clear how to construct a one parameter family by changing the order of commutative operations. More generally, when constructing a one parameter family interpolating between a geodesic γ along which all of $N_1, N_2, \dots N_m$ are discarded, and a geodesic for which none of the submulticurves isotopic to n are discarded, arcs along which surgeries are performed need to be modified to make sense of this.

Suppose γ is the geodesic segment with all the submulticurves isotopic to n discarded as soon as possible. Let γ^1 be a geodesic segment for which all the $\{N_i\}$ but N_1 are discarded, and let γ_{k+1}^1 be the first vertex of γ^1 that does not coincide with γ_{k+1}. Let $\{c_j^{k+1}\}$ be the set of arcs along which γ_k^1 is surgered to obtain γ_{k+1}^1. The set $\{c_j^{k+1}\}$ is obtained by modifying

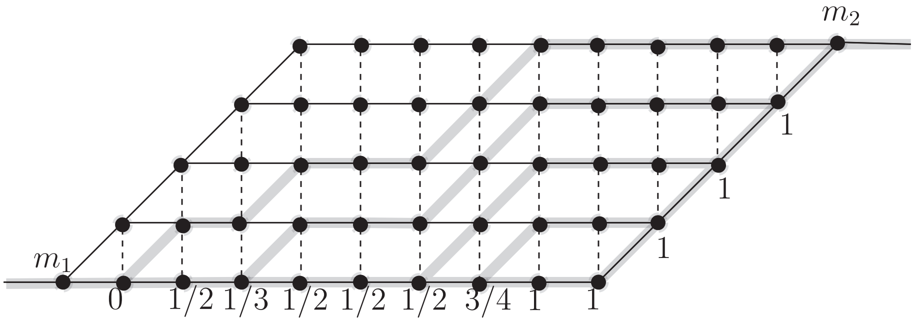

FIG. 9. A one parameter subfamily of the one parameter family from Example 3 is shown in grey. The numbers are the values of the scalar fields on the vertices.

$\{a_j^{k+1}\}$ as follows: if a_j^{k+1} is to the right of γ_k, whenever a_j^{k+1} intersects N_1, replace a_j^{k+1} with the intersection of a_j^{k+1} with the subsurface of S to the right of N_1. Since N_1 is null homologous, this is necessarily a set of horizontal arcs. If a_j^{k+1} does not intersect N_1, leave it unchanged. When the arc a_j^{k+1} is to the left of γ_k, replace it with the intersection of a_j^{k+1} with the subsurface of S to the left of N_1. Similarly, $\{c_j^{k+2}\}$ is obtained by modifying the set $\{a_j^{k+2}\}$ as follows: if a_j^{k+2} is to the right of γ_{k+1}, replace a_j^{k+2} by the set of arcs $a_j^{k+2} \cap (S \setminus \gamma_{k+1}^1)$ (we are assuming the choice of isotopy class described in Subsection 3.3) to the right of γ_{k+1}^1, etc. The last vertex of γ^1 before m_2, call it γ_l, is constructed by surgering along all arcs a_j^l that were not disjoint from γ_{l-1}^1.

The geodesic segment γ^2 is obtained similarly from γ^1, by not discarding the null homologous multicurve N_2, etc.

Given a one parameter family coming from alternative surgeries or isotopic null homologous submulticurves, Jacobi fields are defined in the same way as for optional surgeries.

Restrictions of Jacobi fields. The *restriction* of a Jacobi field $J(a,\gamma)$ can be defined by restricting $J(a,\gamma)$ to a one parameter subfamily $\mathcal{P}_r(a,\gamma) := \gamma, \gamma_r^1, \gamma_r^2, \dots$ of the original one parameter family $\mathcal{P}(a,\gamma) := \gamma, \gamma^1, \gamma^2, \dots$ to which $J(a,\gamma)$ is tangent. An example is shown in Figure 9. In this example γ_r^4 is constructed by surgering along arcs to the left or to the right in a random way. A restriction of the Jacobi field tangent to the one parameter family described in the example interpolates between γ and γ_r^4.

Taking a restriction of a Jacobi field is equivalent to multiplying the magnitude of the vector field $J(a,\gamma)$ by a scalar field $\phi : \{\gamma_i\} \to \mathbb{Q} \cap [0,1]$ defined on the vertices of γ. It is assumed that $0 \leq \phi \leq 1$ is such that $\phi J(a,\gamma)$ determines a valid one parameter family, i.e. $\phi J(a,\gamma)$ is "tangent" to a one parameter family $\mathcal{P}_r(a,\gamma)$ as in the definition of Jacobi field.

REMARK. The definitions of one parameter families are symmetric in m_1 and m_2, but the directions of the Jacobi fields reverse when m_1 and m_2 are interchanged. To understand why this is so, note that surgering along a horizontal arc has an inverse. When m_1 and m_2 are interchanged, this has the effect of exchanging a surgery with its inverse. It follows that the same definition of one parameter family corresponding to the optional surgery s_a, when applied to $m_2, \gamma_j^n, \gamma_{j-1}^n, \dots, m_1$ in place of $m_2, \gamma_j, \gamma_{j-1}, \dots, m_1$, and s_a^{-1} in place of s_a, gives the same family. Exactly the same is true for Jacobi fields arising in other ways. Figure 4 and subsequent Figures were drawn in such a way as to highlight this symmetry.

3.4. Linear Combinations of Jacobi Fields. We would like to be able to describe all geodesics connecting m_1 and m_2 by taking linear combinations of Jacobi fields. To do this, it is necessary to make sure that the linear combination determines a valid set of deformations within one parameter families. There are constraints to check, and it is necessary to make sense of what it means to add Jacobi fields representing surgeries that are not independent.

The constraints are that edges can only connect vertices representing disjoint multicurves whose difference is an embedded, consistently oriented subsurface of S.

Linear combinations of Jacobi fields. The sum of two Jacobi fields $J(a,\gamma)$ and $J(b,\delta)$, where defined, should be thought of as a recipe for moving within two one parameter families. First, $J(a,\gamma)$ determines a deformation of the geodesic γ within a one parameter family to obtain a geodesic γ^k. When $\gamma^k = \delta$, the second Jacobi field gives a recipe for a further deformation within a one parameter family of γ^k. Subtraction of a Jacobi field J is defined as the inverse of addition, i.e. a deformation within a one parameter family in the direction opposite to that determined by J. A more general linear combination of Jacobi fields is the sum or difference of restrictions of Jacobi fields. By a slight abuse of notation, the linear combination of $J(a,\gamma)$ and $J(b,\delta)$ will be called a *linear combination of two Jacobi fields along* γ; $J(b,\delta)$ can often be thought of as a Jacobi field along γ that has been parallel transported through a one parameter family.

Linear combinations of Jacobi fields do not necessarily represent Jacobi fields, because there may not be one parameter families to which the linear combination is tangent.

As an example of a linear combination, let $J(n_1,\gamma)$ and $J(n_2,\gamma)$ be Jacobi fields that arise from discarding non peripheral null homologous multicurves in the isotopy classes n_1 and n_2, respectively. Suppose also γ_i is in the intersection of the support of $J(n_1,\gamma)$ and $J(n_2,\gamma)$, and the subsurface of S bounded by a multicurve in the isotopy class $n_1 - n_2$ is disjoint from the subsurface of S bounded by $\gamma_{i+1} - \gamma_i$. Then $J(n_1,\gamma)$ is a linear combination of $J(n_1 - n_2,\gamma)$ and $J(n_2,\gamma)$.

4. Proof of Theorem 1

To start off with, it will be shown that the Jacobi fields determine the entire space of geodesic segments in some sense. After this, Theorem 1 will be proven.

Let $\mathcal{J}(m_1,m_2)$ be the set of Jacobi fields along geodesic segments connecting m_1 and m_2. As we have seen, $\mathcal{J}(m_1,m_2)$ has the additional structure that linear combinations of some elements are defined.

The subspace of geodesic segments spanned by Jacobi fields along γ. Given two geodesic segments connecting m_1 and m_2, call them δ and γ, δ will be said to be *in the span of* the Jacobi fields along γ if it is possible to find a linear combination of Jacobi fields in $\mathcal{J}(m_1,m_2)$, as defined in Subsection 3.4, that determines a deformation of γ into δ through one parameter families. The *dimension of the space of Jacobi fields along* γ is the smallest possible number of elements of $\mathcal{J}(m_1,m_2)$ needed to span the set of Jacobi fields along γ. The *dimension* of the space of geodesic segments in $C(S,\alpha)$ connecting the vertices m_1 to m_2 is the largest possible dimension of the space of Jacobi fields along a geodesic segment connecting m_1 and m_2.

Theorem 6. *Any geodesic segment connecting m_1 to m_2 is in the span of the Jacobi fields along γ, where γ is a geodesic segment connecting m_1 to m_2 in $C(S, \alpha)$.*

Proof. The geodesic segment γ can be chosen to be the unique middle path in the family of geodesic segments connecting m_1 to m_2. Let δ be the geodesic segment $m_1, \delta_1, \delta_2, \ldots, m_2$. It is sufficient to find a linear combination of Jacobi fields that determines $\gamma - \delta$.

If for every i, δ_{i+1} is constructed by surgering δ_i along arcs on the boundary of S_{imin} and discarding the null homologous multicurves ∂S_{imin}, there is a Jacobi field coming from alternative surgeries that represent the difference of the two geodesic segments. Similarly, whenever for each i, δ_{i+1} could be constructed by surgering along a set of arcs whose endpoints are all assigned the same value of the overlap function, as in Example 3; a surgery of this type is a surgery along the arcs on the boundary of S_{kmin} or S_{kmax} for some k.

The statement of the theorem is also clear when it is possible to reduce to one of these previous cases by subtracting Jacobi fields representing optional surgeries or by adding/subtracting Jacobi fields that represent discarding null homologous submulticurves. The strategy of the rest of this proof is to do just this.

When each δ_{i+1} is constructed from δ_i by discarding null homologous submulticurves not on ∂S_{imax} or ∂S_{imin}, it is possible to add/subtract a Jacobi field to δ to obtain a geodesic segment that does not do this.

Recall that an optional surgery s_a on δ_l may not not define a one parameter family. As discussed in Section 3.3, this happens when s_a is not equivalent to a surgery performed on any of the multicurves representing vertices $\{\delta_i\}, l < i$. For the following special case it will be explained how to find a linear combination of Jacobi fields that take δ to a geodesic for which s_a *does* determine a one parameter family.

Suppose, for some $l < k$, $\{a^k_j\}$ can be chosen such that

- for each j the endpoints of the arcs a^k_j have the same value f of the overlap function of $m_2 - m_1$, and
- f is the value of the overlap function of $m_2 - m_1$ on the endpoints of a.

This special case occurs, for example, when all surgeries except s_a are along arcs on the boundary of S_{imax} or all surgeries except s_a are along arcs on the boundary of S_{imin}. There is a Jacobi field J coming from an alternative surgery that replaces surgeries along the arcs $\{a^k_j\}$ with surgeries on δ_m, $k \leq m$, along arcs $\{c^{m+1}_j\}$ with the same end points as $\{a^k_j\}$, but on the other side of δ_m. We saw an example of a one parameter family that does this in Example 3.

Assuming the special case, moving δ in the direction of J, a geodesic segment is obtained along which s_a determines a one parameter family. Subtract the corresponding Jacobi field to obtain a geodesic segment with one fewer optional surgeries than δ. The resulting geodesic is then closer to the type of geodesic we are trying to obtain.

Now if the previous special case does not occur, and s_a is the only optional surgery along δ, the remainder of this proof, applied to the geodesic segment connecting δ_{l+1} to m_2, shows how to reduce to the special case from the previous paragraph. If there is more than one optional surgery, let s_a be the optional surgery performed on δ_l, where δ_l is the last multicurve representing a vertex of δ along which optional surgeries are performed. Whenever two or more optional surgeries are performed on δ_l, the corresponding arcs are necessarily on the same side of δ_l, so the Jacobi field J takes δ to a geodesic segment for which both

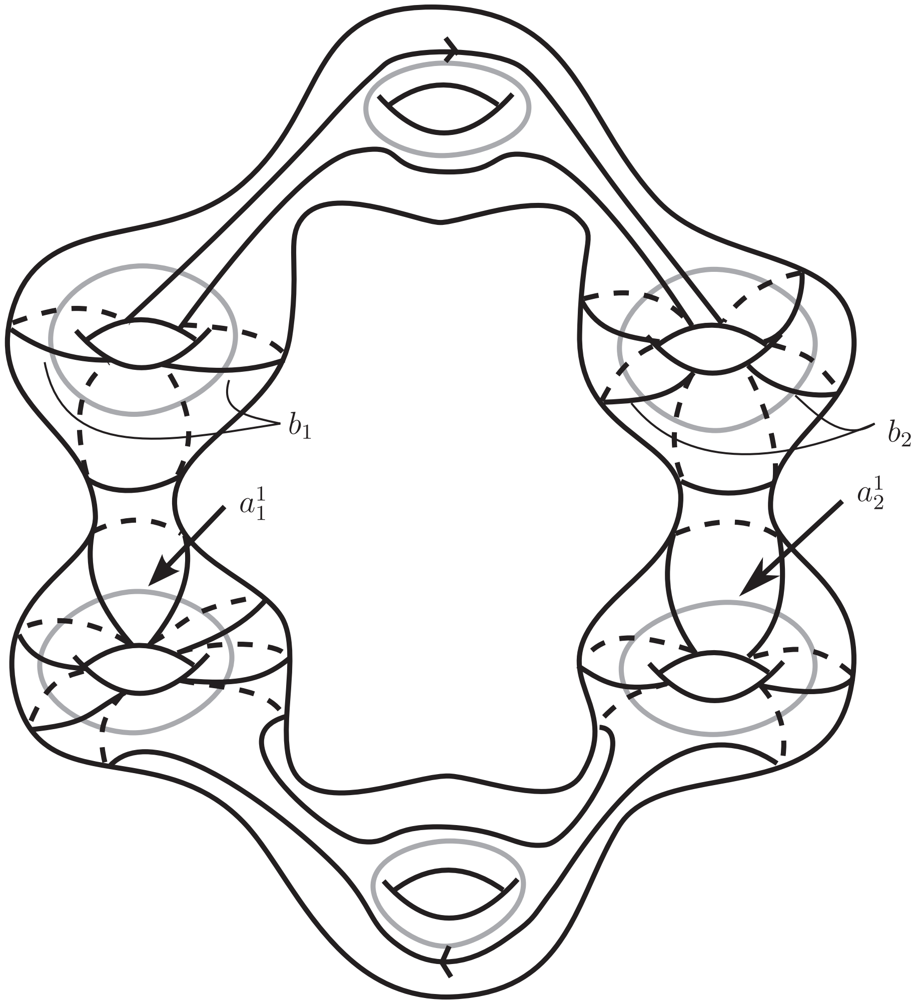

FIG. 10. The multicurve m_1 is black and m_2 is shown in grey.

of the optional surgeries determine one parameter families. After removing the last optional surgery, then the second last optional surgery is removed, etc.

The main difficulty in proving this theorem comes from examples such as that in Figure 10. Suppose δ_1 is constructed by surgering m_1 along the arcs $\{a_1^1\}$ and $\{a_2^1\}$ to the right of m_1. This set of arcs can not be chosen such that their endpoints have the same value of the overlap function, and none of the corresponding surgeries are optional. Call such sets of surgeries *diagonal*.

Alternative surgeries and partitions of multicurves. Recall that a condition for δ to be a geodesic is that $\delta(\delta_{i-1}, m_2) = \delta(\delta_i, m_2) + 1$, where $\delta(*, *)$ denotes the homological distance. Let $\{a_j^1\}$ be the set of arcs along which m_1 is surgered to obtain δ_1. A connected subarc of m_1 whose intersections with ∂S_{max} and ∂S_{min} are nonempty must have a connected subarc with algebraic intersection number $\delta(m_1, m_2)$ with $m_2 - m_1$. If this subarc is contained in a connected component of $m_1 \setminus \{\partial a_j^1\}$, it will remain after surgering along $\{a_j^1\}$ and still have algebraic intersection number $\delta(m_1, m_2)$ with $m_2 - m_1$. This contradicts the requirement that

$\delta(m_1, m_2) = \delta(\delta_1, m_2) + 1$. Similarly, if there were a connected component of $m_2 \setminus \{\partial a^1_j\}$ with arcs on the boundary of both S_{max} and S_{min}, the condition $\delta(m_1, m_2) = \delta(\delta_1, m_2) + 1$ could not be fulfilled. It follows that the endpoints of the arcs $\{a^1_j\}$ separate S_{max} from S_{min} on both multicurves m_1 and m_2. The multicurve δ_1 is therefore a union of two disjoint multicurves; δ_{1+}, which has arcs on the boundary of S_{max}, and δ_{1-}, which has arcs on the boundary of S_{min}. Further diagonal surgeries along arcs, each with endpoints on δ_{1+} or δ_{1-} (but not both) to obtain the multicurve δ_2 will also clearly preserve the decomposition.

To keep the notation simple, suppose to start off with that δ is constructed by performing as many diagonal surgeries as possible, as follows: keep performing diagonal surgeries to construct consecutive vertices along δ until a value of i, call it i^*, is reached such that no diagonal surgery on δ_{i^*} can be used to construct the next vertex along δ. This happens when the maximum of the overlap function has been brought so low that δ_{i^*-} has an arc on the boundary of S_{i^*max} and the minimum of the overlap function brought so high that δ_{i^*+} has an arc on the boundary of S_{i^*min}.

The surgeries along arcs with endpoints on different multicurves commute, and hence determine one parameter families corresponding to alternative surgeries. Therefore, by moving δ through one parameter families to the geodesic segment η, it is possible to assume without loss of generality that the surgeries on the "−" multicurves were performed before those on the "+" multicurves. Similarly, if diagonal surgeries on δ are interspersed with other surgeries, η is chosen such that the diagonal surgeries were all performed first. Since for $i^* < i$ it is not possible to perform any more diagonal surgeries, it can be assumed without loss of generality that for $i^* < i$, η_{i+1} is constructed by surgering η_i along the arcs on the boundary of S_{imin}. The reason this can be assumed without loss of generality is that if this is not the case, we have already discussed how to move η through one parameter families to achieve this.

Let η_k, $k \leq i^*$ be the first vertex of η at which we start perfoming diagonal surgeries on the "+" multicurves. By construction, there is an arc or arcs of $m_2 \cap (S \setminus \eta_{k+})$ on the boundary of S_{kmin}. Arcs of $m_2 \cap (S \setminus \eta_{k-})$ on the boundary of S_{kmin} represent optional surgeries and determine one parameter families over η. Move η into these one parameter families to obtain η^1. On η^1, the surgeries on η_{k+} along arcs not on the boundary of S_{kmin} become optional surgeries that determine a one parameter family. Remove these optional surgeries by moving through the corresponding one parameter families to get a geodesic segment μ, where μ_{k+1} was constructed from μ_k by surgering along arcs on the boundary of S_{kmin}. Similarly for $\mu_{k+2}, \mu_{k+3}, \ldots, \mu_{i^*}$.

Except for the diagonal surgeries on μ_{i-} for $i < k$, each μ_{i+1} is constructed by surgering μ_i along arcs on ∂S_{imin}.

Now starting with μ_k, replace the surgeries along arcs of $m_2 \cap (S \setminus \mu_k)$ on the boundary of S_{kmin} with the surgeries along the arcs of $m_2 \cap (S \setminus \mu_k)$ on the boundary of S_{kmax}. Do the same with μ_{k+1}, μ_{k+2} etc. until μ_j is reached, $k < j$, where μ_{j-} has an arc on the boundary of S_{jmax}. This is done by moving through one parameter families corresponding to alternative surgeries to get to the geodesic segment ν. Along ν, for $i < k$, the surgeries used to construct the first k multicurves are diagonal surgeries on ν_{i-}. These surgeries therefore commute with the surgeries for constructing the next $j - k$ multicurves, because for $k \leq i \leq j$ these surgeries are all along arcs on ∂S_{imax} and therefore surgeries on ν_{i+}. Again, moving through one parameter families, it is therefore possible to exchange the order, to obtain a geodesic segment ω. Along ω, let ω_l be the first vertex at which we start surgering along the "−" multicurves. By construction, ω_{l-} has an arc on the boundary of S_{imax}. Therefore, the same argument as before shows that it is possible to get rid of the remaining diagonal surgeries by moving through one parameter families.

If δ is constructed by performing fewer than the maximum possible number of diagonal surgeries, by moving through one parameter families, the same arguments show that there are one parameter families through which δ can be moved to reach the middle path.

For any i, δ_{i+1} is constructed from δ_i by a combination of the following:

- discarding a null homologous submulticurve,
- performing one or more optional surgeries,
- performing diagonal surgeries,
- performing surgeries that are not diagonal.

We have therefore covered all the different types of surgeries or ways of discarding null homologous multicurves that might be used to construct a geodesic path, and shown that there exist linear combinations of Jacobi fields that take vertices on all these geodesics to corresponding vertices on the middle path. □

Critical Level. Recall that a geodesic segment in $\mathcal{C}(S,\alpha)$ with endpoints m_1 and m_2 is an indexed set of vertices $m_1, \gamma_1, \gamma_2, \dots, \gamma_j, m_2$ for which $d(\gamma_i, \gamma_{i+1}) = d(\gamma_j, m_2) = d(m_1, \gamma_1)$. Consider the set $\mathcal{J}(\gamma)$ of Jacobi fields along γ. The index i is a *critical level* if γ_i is the first or last vertex in the support of an element $J(a,\gamma)$ of $\mathcal{J}(\gamma)$, where $J(a,\gamma)$ is not the restriction of another Jacobi field.

The index i could be a critical level if, for example, the vertex after γ_{i-1} could not have been constructed by surgering along a set of arcs of the form $\{v_j \circ a_j^{i-1} \circ w_j\}$, or when the number of arcs in the homotopy class with representative $v_j \circ a_j^{i-1} \circ w_j$ is not the same as the number of arcs in the homotopy class with representative a_j^{i-1} for some j.

Remark. There are two possible ways in which the dimension of the space of geodesic segments could have been defined. Firstly, in terms of the maximum possible number of Jacobi fields along a geodesic segment as done here, and secondly, in terms of the maximum number of Jacobi fields needed in a linear combination representing the difference of two geodesic segments. Analysing the proof of Theorem 6 carefully shows that, assuming Theorem 1, both are finite. This is because it is possible to move δ through a finite number of one parameter families to a geodesic segment ω for which the following is true: For all i, ω_{i+1} is constructed from ω_i by the obvious continuation of the construction of ω_i from ω_{i-1} unless ω_i is a critical level. Then the deformations that take a vertex ω_{i+1} to its target vertex γ_{i+1} are the obvious continuations (i.e. deformations within the same one parameter family) of the deformations needed to take ω_i to its target vertex γ_i, unless a critical level is reached.

It follows from the remark that when geodesic segments with the same endpoints do not stay close, there will necessarily be some Jacobi field with large magnitude.

We now begin the proof of Theorem 1.

Proof. The number of homotopy classes of arcs of $m_2 \cap (S \setminus m_1)$ is bounded. For example, in [6], Lemma 11, the sharp bound $-3\chi(S)$ was obtained. Given an arc a of $m_2 \cap (S \setminus \gamma_i)$, let v, w be the arcs of $m_2 \cap (S \smallsetminus \gamma_i)$ on either side of a. To see how the number of homotopy classes of horizontal arcs bounds the dimension of the space of Jacobi fields, first note that surgeries along homotopic arcs are equivalent. Secondly, if γ_i is surgered along a set of horizontal arcs $\{a_j^{i+1}\}$ containing the arc a, the set $\{a_j^{i+1}\}$ does not also contain the arc $v \circ a \circ w$ because

- if a has both endpoints on a curve c in γ_i such that γ has more than one curve homotopic to c, then $\gamma_{i+1} - \gamma_i$ could not be the boundary of an embedded, oriented subsurface of S. If it were, then $\delta(\gamma_{i+1}, \gamma_i) > 1$, contradicting Corollary 2.
- if a has both endpoints on the null homologous curve $N(\{a_j^{i+1}\})$ discarded after surgering along $\{a_j^{i+1}\}$, then the null homologous submulticurve obtained by surgering $N(\{a_j^{i+1}\})$ along $v \circ a \circ w$ is discarded. It does not make any difference to the path if we surger $N(\{a_j^{i+1}\})$ along $v \circ a \circ w$ before discarding it or not. It can therefore be assumed without loss of generality that this does not happen.
- in all other cases, surgering along $v \circ a \circ w$ would mean that γ_{i+1} intersects γ_i. This is because the arc $v \circ a \circ w$ crosses over γ_i, and it is not possible to remove these points of intersection by an isotopy as in the first case, or by discarding a null homologous multicurve as in the second case.

It follows that there are at most $-3\chi(S)$ Jacobi fields with support on any given vertex. However, we are trying to prove something a bit stronger than that, namely that the total number of Jacobi fields along a geodesic is bounded.

Local extrema of the overlap function can not ever be created as i increases; surgering along a horizontal arc of $m_2 \cap (S \smallsetminus \gamma_i)$ to the right of γ_i decreases a local maximum along m_2, and surgering γ_i along a horizontal arc to the left of γ_i increases a local minimum along m_2. A saddle of the overlap function is a local extremum along m_2, so for the same reason, the number of saddles can not increase either. However, this does not immediately give a bound on the number of homotopy classes of horizontal arcs of $m_2 \cap (S \smallsetminus \gamma_i)$ because not all saddles or local extrema determine independent surgeries. Many of them might have homotopic arcs on their boundaries. In Figure 11 is an example of how the number of *homotopy classes* of horizontal arcs can increase.

Splitting and Killing homotopy classes. For a given arc a in the set $\{a_j^1\}$, suppose $v_1 \circ a \circ w_1$ is an arc in the set $\{a_j^2\}$, and $v_2 \circ v_1 \circ a \circ w_1 \circ w_2$ an arc in $\{a_j^3\}$, etc. For large enough n, one or both of the following two things will happen: there are two or more homotopy classes of arcs $v_n' \circ \ldots \circ a \circ w_1 \circ \ldots \circ w_n$ and $v_n'' \circ \ldots \circ a \circ w_1 \circ \ldots \circ w_n$ or $v_n \circ \ldots \circ a \circ w_1 \circ \ldots \circ w_n'$ and $v_n \circ \ldots \circ a \circ w_1 \circ \ldots \circ w_n''$; this will be called *splitting* the homotopy class a. The other possibility is that $v_n \circ \ldots \circ a \circ w_1 \circ \ldots \circ w_n$ is a vertical arc, but $v_{n-1} \circ \ldots \circ a \circ w_1 \circ \ldots \circ w_{n-1}$ was not. This will be called *killing* the homotopy class $v_{n-1} \circ \ldots \circ a \circ w_1 \circ \ldots \circ w_{n-1}$. A homotopy class is killed when one, but not both, of v_n or w_n is a horizontal arc.

Once a homotopy class of horizontal arcs has been killed, the resulting homotopy classes of vertical arcs can become v_is and w_is for another horizontal arc. The hexagons, octagons etc. that split a into homotopy classes can then cause another homotopy class of horizontal arcs to be split.

We have seen that surgering the multicurve γ_i along horizontal arcs of m_2 can not create local extrema of the overlap function along m_2, and up to homotopy, there were no more

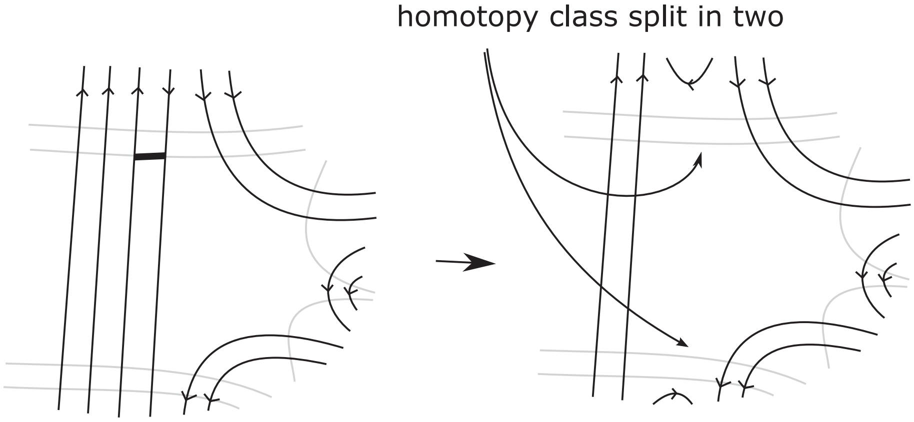

FIG. 11. The multicurve m_2 is shown in grey, and m_1 in black. After surgering along the horizontal arc of $m_2 \cap (S \setminus m_1)$ indicated by the thick black line, the number of homotopy classes of horizontal arcs increases.

than $-3\chi(S)$ arcs of $m_2 \cap (S \setminus m_1)$ representing local extrema on m_2. Therfore, since there were at most $-3\chi(S)$ homotopy classes of horizontal arcs to start off with, and each can be split into at most $-3\chi(S)$ homotopy classes of horizontal arcs, this gives a bound of $9\chi(S)^2$ for the number of Jacobi fields along γ coming from optional surgeries.

Some of the optional surgeries, when grouped together, might determine Jacobi fields coming from alternative surgeries. Also, a given homotopy class of horizontal arcs might determine a surgery that is performed as a component of more than one alternative surgery.

There can be no more than $-2\chi(S)-2$ Jacobi fields from alternative surgeries with support on the vertex γ_1 of γ. This follows from the observation in the proof of Theorem 6, in the paragraph titled "Alternative surgeries and partitions of multicurves". We saw that an alternative surgery determines a null homologous multicurve, call it ∂S_+, that partitions γ_1 into two multicurves, γ_{1+} and γ_{1-}. For surgeries along arcs on the boundary of S_{max} or S_{min}, ∂S_+ could be contractible, giving a trivial partition. Suppose $\{a_j^1\}$ determines an alternative surgery at the vertex γ_1 of γ, cutting off a null homologous curve n. If $\{v_j \circ a_j^1 \circ w_j\}$ also represent alternative surgeries, the null homologous curve n_2 cut off will be to one side of, i.e. disjoint from, n; otherwise surgering along $\{v_j \circ a_j^1 \circ w_j\}$ could not give a vertex distance one from γ_1. Similarly, n_3, the next null homologous curve cut off by surgeries along arcs of the form $\{v_j^2 \circ v_j \circ a_j^1 \circ w_j \circ w_j^2\}$ must be disjoint from n_2 and on the other side of n_2 as n, hence disjoint from n, etc. The factor of 2 in $-2\chi(S) - 2$ comes from the fact that two distinct sets of arcs of $m_2 \cap (S \setminus \gamma_1)$ might cut off the same separating multicurve; one set of arcs on the left of γ_1 and the other on the right. An example of this is the sets of arcs labelled $\{a_1^1, a_2^1\}$ and $\{b_1, b_2\}$ in Figure 10.

An upper bound on the number of isotopy classes of null homologous submulticurves giving linearly independent Jacobi fields is half the number of Jacobi fields coming from alternative surgeries. In total, this gives a bound of $9\chi(S)^2 - 3\chi(S) - 3$. □

REMARK. The bound in the previous proof is clearly not sharp. However, to get a considerably better bound, it would seem that a much more detailed argument would be needed; the details of which are more tedious than illuminating.

5. Sublevel Projections

Subsurface projections were defined in [10] in order to be able to break the curve complex down into simpler pieces, thought of as curve or arc complexes of subsurfaces. The nested structure arising from the subsurface projections were used to describe families of quasigeodesics called hierarchy paths, and to show how these families of quasigeodesics are controlled by the subsurface projections of their endpoints.

In this section, the notion of sublevel projections are defined, so-named because there are some very strong parallels with subsurface projections. Informally, critical levels are used to partion a geodesic into subintervals that are as rigid as possible and behave almost independently of each other.

Let $m_1, \gamma_1, \ldots, m_2$ be the middle path connecting m_1 and m_2. Given two integers $l_1 < l_2$ in the range of the overlap function of $m_2 - m_1$, the *sublevel projection* of m_1 and m_2 between the levels l_1 and l_2, $\Pi_{l_1}^{l_2}(m_1, m_2)$, is the pair of homologous multicurves $(\gamma_{l_1+1}, \gamma_{l_2})$.

The sublevel projection of m_1 and m_2 between the levels l_1 and l_2 is similar to a subsurface projection to $S_{l_1+1\leq f\leq l_2}$, in the sense that γ_{l_1} and γ_{l_2} represent vertices as close as possible to m_1 and m_2, respectively, given that they only intersect within the subsurface $S_{l_1+1\leq f\leq S_{l_2}}$. It follows from Theorem 9 in [6] that this definition is symmetric in m_1 and m_2.

Distance Formula. Consider the finite number of sublevel projections of the form $\Pi_i := \Pi_{l_i}^{l_{i+1}}(m_1, m_2)$, where l_i and l_{i+1} are critical levels. Any collection of surgeries performed on the multicurve γ_i to construct a multicurve γ_{i+1} with $d(\gamma_{i+1}, \gamma_n) = d(\gamma_i, \gamma_n) - 1$ necessarily decreases the distance between γ_i and γ_n in one of the sublevel projections Π_i. A distance formula analogous to the distance formula from [10] Section 6, with a uniform bound on the number of sublevel projections follows immediately from the construction and Corollary 2. In this way, families of tight paths in $C(S, \alpha)$ are even more rigidly controlled by the sublevel projections of their endpoints than is the case in the marking graph for hierarchy paths under subsurface projections, [10].

Acknowledgements. This work was funded by a MOE AcRF-Tier 2 WBS grant Number R-146-000-143-112. The author is grateful to M. Korkmaz and an anonymous reviewer whose comments and attention to detail greatly improved the readability of this paper.

References

[1] A. Hatcher and D. Margalit: *Generating the Torelli Group*, Enseign. Math. **58** (2012), 165–188.

[2] M. Bestvina, K. Bux and D. Margalit: *The dimension of the Torelli group*, J. Amer. Math. Soc. **23** (2010), 61–105.

[3] B. Bowditch: *Tight geodesics in the curve complex*, Invent. Math. **171** (2008), 281–300.

[4] M. do Carmo: Riemannian Geometry, volume 1, Birkhäuser, 1992.

[5] A. Hatcher: *The cyclic cycle complex of a surface*, arXiv:0806.0326v1, 2008.

[6] I. Irmer: *The geometry of the homology curve complex*, J. Topol. Anal. **4** (2012), 335.

[7] I. Irmer: *The Chillingworth class is a signed stable length*, Algebr. Geom. Topol. **15** (2015), 1863.

[8] J. Schultens: *The Kakimizu complex of a surface*, arXiv:1401.2111v3, 2014.

[9] H. Masur and Y. Minsky: *Geometry of the complex of curves I: Hyperbolicity*, Invent. Math. **138** (1999), 103–149.

[10] H. Masur and Y. Minsky: *Geometry of the complex of curves II: Hierarchical Structure*, Geom. Funct. Anal. **10** (2000), 902–974.
[11] A. Putman: *Cutting and Pasting in the Torelli group*, Geom. Topol. **11** (2007), 829–865.
[12] W. Thurston: *A norm for the homology of 3-manifolds*, Mem. Amer. Math. Soc. **59** (1986), 99.

Department of Mathematics and Statistics
University of Melbourne
Parkville, VIC 3010
Australia
e-mail: ingrid.irmer@unimelb.edu.au